国之重器出版工程

网络强国建设

学术中国·大数据

Big Data
Processing Platform

大数据处理平台

宋杰 著

中国工信出版集团

人民邮电出版社
POSTS & TELECOM PRESS

大数据处理平台

Big Data
Processing Platform

宋杰 著

人民邮电出版社
北京

图书在版编目（CIP）数据

大数据处理平台 / 宋杰著. -- 北京 : 人民邮电出版社, 2017.12
（学术中国. 大数据）
国之重器出版工程
ISBN 978-7-115-46689-1

Ⅰ. ①大… Ⅱ. ①宋… Ⅲ. ①数据处理 Ⅳ. ①TP274

中国版本图书馆CIP数据核字(2017)第261361号

内 容 提 要

本书从数据查询、数据分析和迭代计算平台3个方面对大数据处理平台的体系结构、基本原理、主流技术、国内外研究进展和成果进行了全面、深入的阐述，对大数据实时处理平台的架构和核心技术进行了展望。企业技术人员可参考本书选择合适的技术构建大数据处理平台或对现有平台进行优化；高校及研究院所的科研人员可参考本书了解大数据管理的基本原理和现有研究成果；高校学生可通过学习本书全面了解大数据处理平台。同时，本书也适用于对大数据技术有浓厚兴趣的读者。

◆ 著　　宋　杰
责任编辑　吴娜达
责任印制　杨林杰
◆ 人民邮电出版社出版发行　　北京市丰台区成寿寺路11号
邮编　100164　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
固安县铭成印刷有限公司印刷
◆ 开本：720×1000　1/16
印张：14.75　　2017年12月第1版
字数：220千字　　2017年12月河北第1次印刷

定价：89.00元

读者服务热线：(010)81055488　印装质量热线：(010)81055316
反盗版热线：(010)81055315
广告经营许可证：京东工商广登字20170147号

《国之重器出版工程》
编辑委员会

苏哲子　中国工程院院士

李伯虎　中国工程院院士

李应红　中国科学院院士

李新亚　国家制造强国建设战略咨询委员会委员、中国机械工业联合会副会长

杨德森　中国工程院院士

张宏科　北京交通大学下一代互联网互联设备国家工程实验室主任

陆建勋　中国工程院院士

陆燕荪　国家制造强国建设战略咨询委员会委员、原机械工业部副部长

陈一坚　中国工程院院士

陈懋章　中国工程院院士

金东寒　中国工程院院士

周立伟　中国工程院院士

郑纬民　中国计算机学会原理事长

郑建华　中国科学院院士

屈贤明　国家制造强国建设战略咨询委员会委员、工业和信息化部智能制造专家咨询委员会副主任

项昌乐　“长江学者奖励计划”特聘教授，中国科协书记处书记，北京理工大学党委副书记、副校长

柳百成　中国工程院院士

闻雪友　中国工程院院士

徐德民　中国工程院院士

唐长红　中国工程院院士

黄卫东　“长江学者奖励计划”特聘教授

黄先祥　中国工程院院士

黄　维　中国科学院院士、西北工业大学常务副校长

董景辰　工业和信息化部智能制造专家咨询委员会委员

焦宗夏　“长江学者奖励计划”特聘教授

《学术中国·大数据》丛书
编辑委员会

丛书总序

大数据、人工智能、云计算、物联网、移动互联网和产业互联网等成为新一代信息技术的特征，其中大数据与上述技术和应用都有密切关系。大数据来自于移动互联网、产业互联网和物联网等，其存储需要云计算，其挖掘依靠人工智能，而人工智能也有赖于大数据的支撑，大数据是产业互联网的重要基础。大数据不仅可以用于社会的精细化管理，更好地服务民生，大数据产业也将形成信息产业新的分支，其间接的产业影响将更大。可以说，大数据是数字经济的重要支柱。

很多国家都将大数据作为新时期的国家发展战略。2015 年，国务院印发大数据发展的首个权威性、系统性文件《促进大数据发展行动纲要》，2016 年国家发展和改革委员会批复了 13 个大数据领域的国家工程实验室，我国一些省市也纷纷制定大数据发展战略与规划。当前，我国在大数据共享开放、大数据资源开发、大数据技术研发、大数据挖掘应用、大数据产业培育、大数据安全管理、大数据人才培养和大数据法规研究等方面全面部署，为我国实现供给侧结构性改革，促进产业升级和转型，提升国家竞争力，争取在国际领域的话语权和实现跨越式发展起到了不可或缺的作用。

然而，我国的大数据发展也面临一些亟待解决的问题，例如基础研究薄弱、创新能力不强、产业链条缺口、数据资源封闭、法律法规滞后、数据安全不力、数据人才短缺和数据设施布局不合理及利用率不高等。为了使我国的大数据应用与产业可持续健康发展，需要多管齐下，其中普及大数据科学是重要的一环。为此，《学术中国·大数据》丛书编委会组织多个大数据领域优秀的研究团队的专家，基于国家

“973”计划、“863”计划、国家自然科学基金、国家重点研究计划等科研项目的创新研究成果和国内外大数据应用的成功实践，编写了这套丛书，内容涵盖大数据存储、数据管理、数据挖掘、分析平台、优化算法等核心技术领域。

本丛书的出版对传播大数据科学知识、推动大数据的学术探讨、鼓励大数据领域的产学研用协同创新、促进大数据标准化研究、加快大数据核心技术研发、培训大数据技术人才、引导大数据应用与产业化发展以及完善大数据有关的制度建设，都将起到积极作用。

2017 年 12 月

前 言

如何从海量数据中有效获取信息，以进行分析和决策是大数据的核心问题，也是21世纪各行各业均面临的重要问题。解决这一核心问题需要大数据处理平台的支持。大数据处理平台是一种“计算平台”，计算平台泛指支持算法执行的硬件系统、操作系统和运行库，那么大数据处理平台则泛指可以支持大数据处理算法执行的平台。大数据处理平台采用集群作为硬件系统，分布式计算框架作为中间件系统。以Hadoop HDFS为代表的分布式文件系统和以MapReduce为代表的分布式并行编程模型在学术界和产业界最为流行，并以此引出完善的Hadoop生态圈；另一个则是围绕并行框架Spark的生态系统，如Spark Streaming和Shark。以这些开源技术为支撑的大数据处理平台广泛地应用于社交网络、科学数据分析、传感数据处理、医疗和电子商务平台中。

典型的大数据应用可以分为OLTP、OLAP和图计算3类，因此，从数据处理平台角度，需要提供数据查询、统计分析和迭代计算支持。本书围绕这3个典型数据处理方式，首先介绍大数据处理平台的体系结构，并简述体系结构每部分的主流技术；随后重点介绍大数据处理平台实现数据查询、统计分析和迭代计算的基本原理、研究进展；每部分还包括项目组近年来的研究成果。此外，本书还展望了大数据实时处理平台的架构和核心技术。本书介绍的理论和技术均集中于中间件层，以“学术研究”和“系统实现”相结合的角度论述，使得读者能够更加深入地理解大数据处理平台的核心技术和学术前沿，帮助读者更加有效地构建处理平台，或对已有的大数据处理平台进行改进，开展大数据存储和管理领域的相关研究。

本书是“学术中国·大数据”系列丛书之一。书中研究成果为国家自然科学基金重点项目“大数据高效能存储与管理方法研究（No.61433008）”的部分建设成果。笔者在数据管理领域已有十多年的研究经验，结合自身的研究经验，从“学术研究”和“系统实现”相结合的角度，对平台进行全面的介绍。书中既有原理，又有学术前沿综述，但不包含使用方法、编程技术、构建步骤等类似工具书的内容。对于大数据相关领域高校师生、研究人员以及大数据处理平台的设计师和架构师有一定的借鉴性。

笔　者

2017 年 4 月于沈阳南湖

目 录

第1章
体系结构

大数据处理平台是一种“计算平台”，计算平台泛指支持算法执行的硬件系统、操作系统和运行库，大数据处理平台泛指可以支持大数据处理算法执行的平台。大数据处理平台采用集群作为硬件系统，分布式计算框架作为中间件系统。本章主要介绍大数据处理平台的体系结构，并简述体系结构每部分的主流技术。

1.1 集群系统

集群系统通常是指构建在计算机集群之上的系统，其将一定数量的计算机连接起来构成分布式系统，作为单独的、统一的计算资源，为上层应用提供计算服务。集群系统对同构或异构的计算资源进行统一管理与调度以及实现分布式协调（Distributed Coordination）。本节主要介绍当前主流的任务管理与调度机制 Hadoop YARN、Apache Mesos 以及分布式协调机制 Apache ZooKeeper。

1.1.1 Hadoop YARN

由于 MapReduce 的 JobTracker/TaskTracker 机制在可扩展性、内存消耗、线程模型、可靠性等方面存在缺陷，且其维护成本很高，因此，Apache 提出了 Hadoop YARN 这种新型的 Hadoop MapReduce 框架。YARN 架构如图 1-1 所示。YARN[1] 提供强大的资源管理功能，将 JobTracker 分解为两个独立的服务：全局的资源管理器（Resource Manager，RM）和应用程序特有的应用程序管理器（Application Master，AM）。YARN 采用 Master-Slave（主—从）架构，主节点部署 RM，负责整个系统的资源管理和分配。每个从节点上均部署节点资源管理器（Node Manager，NM），RM 对各个 NM 上的资源进行统一管理和调度。